

#5



PCT09

ENTERED

RAW SEQUENCE LISTING

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

```

4 <110> APPLICANT: Ruelle, Jean-Louis
6 <120> TITLE OF INVENTION: IMMUNOGENIC COMPOUNDS
9 <130> FILE REFERENCE: BM45378
11 <140> CURRENT APPLICATION NUMBER: 09/914,168
C--> 12 <141> CURRENT FILING DATE: 2002-01-28
14 <150> PRIOR APPLICATION NUMBER: 9904559.3
15 <151> PRIOR FILING DATE: 1999-02-26
17 <150> PRIOR APPLICATION NUMBER: PCT/EP00/01468
18 <151> PRIOR FILING DATE: 2000-02-23
20 <160> NUMBER OF SEQ ID NOS: 6
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 2760
26 <212> TYPE: DNA
27 <213> ORGANISM: Moraxella catarrhalis
29 <400> SEQUENCE: 1
30 atgtcaaagc cgtttttggt tgcaaactgc agtttttatgc ctgtcgcatt ggcggtttat 60
31 ttgcctttga tgacatcgca agcattggca caacaaaata accctgcaaa catcatcaat 120
32 catgtacccg ctcatgacac cgccatcaat caagcaaaagg caggcaatcc gcctgttttg 180
33 ctaacacctg agcagatata agcacgcctt aatgctgctg gactgaatgc taagccccaa 240
34 tcacaagctt tggatgttgt caattttgat gatcaatcgc cgatatctcg tatcgggtgag 300
35 caatcacccc ctttgggttt ggatatgtcg gtcacgaag aaaccacacc gctaagcttg 360
36 gaggaattat ttgctcaaga atctactgag atgggaatca atccaaatga ttatatcca 420
37 gaatatcaag gcgagcaacc taatagttag gtggtgttac caccgacatt agaacctgaa 480
38 aaaccagggt tgatcaagcg tctttatgca cgcctattta atgatgggtg caataagggtg 540
39 cctaggetta aggcaaaatt ttatcaatca tcgcaatcag gcgaaaccag tgcgattggg 600
40 tcatcgcata aaaaaacaga gccttatgca aatatcaaag cagcacttga agacatcacc 660
41 caagagtcag cgatggattt gaatggctct atcccacgcc taaggcaaac tgctttggtg 720
42 gcagcgcgtg ctgtcggtta ttatgatatt gatttatcaa tcataagaaa tagcatcgga 780
43 gaggtggatg tcatcatcca tgatttaggt gaacctgttt atattgatta tcgagcgggtg 840
44 gaggtacgag gtgaagggtg tgatgataaa gcatttacta ccgtggcgga tgaggtgcca 900
45 ttgctgatcg gcgatgtctt tcatcatggc aagtacgaaa ccaaaaaaaaa tctcatcgaa 960
46 aatgccagtg ctgaacatgg atattttgat gggcggttggc tggatcggtc agttgatgta 1020
47 attttgccag ataataccgc tgatgtcagc ttaatttatg atacagggtac gcagtatcgc 1080
48 ttgatgagg tggtattttt taccattgat cctaaaacca atcaattgac aaccgatcca 1140
49 gataagctgc cagttaaacg agaattactt gagcagttac tcaccgttaa catgggagag 1200
50 gcttacaaat tacaggcggg gcgtgcactt tcaaatgatt tgattgccac acggtatttt 1260
51 aatatggtga ataccgagat tgtctttcca gagcgtgaac agatccaaaa cgaccaagtg 1320
52 agctttgagc agtcttcaag tagccgtact gaaccagcac aagttgatga aagcacactt 1380
53 gaacctgtca ttgaaaccgt tgagctaacg gatgggatat taatggatat ttcgcccatc 1440
54 gaatttagtg catctaactt gattcaagac aagctaaatt tgggtggctgc caaggctcgc 1500
55 catttatatg acatgcctga tgataggggtg cttgccatca atcatgatga tggcgtaaat 1560
56 cgctctattt tgggcagaat cagcgatgcc gtatctgcgg ttgcacgtgc tattttacct 1620

```

RAW SEQUENCE LISTING

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

```

57 gatgaatctg aaaatgaggt aatagatttg cccgagcgta ccgcattggc taatcgcaag 1680
58 acccctgctg atgtctatca aagtaaaaaa gtgccgctat atgtctttgt ggcgagtgat 1740
59 aaaccacgag atggtcaaat tggtttgggc tggggatcgg acacaggtag ccgcctagtc 1800
60 acaaaaatttg agcataatth gattaatcgt gatggctatc aagcaggcgc tgagctaaga 1860
61 ctgtctgagg ataaaaaagg ggtcaagtta tatgccacca aaccgcttag ccaccctcta 1920
62 aatgatcagc taagagcaac tttgggttat caacaagaag tttttggtca ctctaccaat 1980
63 ggttttgatt tatccacacg caccctagag catgagatta gccgcagtat tatccaaaat 2040
64 ggtggctgga atcgtactta ttcattgcgt tatcgtcttg ataagcttaa aaccaagca 2100
65 cccctgaaa catggcagga tttaccagtg gattttgtca atggtaagcc aagccaagag 2160
66 gcgttattgg cagggtgtgc tgtgcataaa acggttgtag ataatttggg taatccgatg 2220
67 cgtggctatc gtcagcgata ttcttttagag gttggctcaa gcggtttggg atcggatgct 2280
68 aatatggcta ttgctcgagc tggatttagt ggcgtgtata gttttgggga taatgcttat 2340
69 ggcagcaatc gtgcccatac gatgactggg gccatacaag caggatacat ttggtcggat 2400
70 aattttaatc atgtgccata tctgttgcgt ttttttgctg gtggcgacca aagtattcgt 2460
71 ggatatgcac atgacagttt atcacctata tcagataagg gttatctgac aggcgggtcaa 2520
72 gtattggcgg ttggtacagc tgaatataat tatgaattta tgaaagattt gcgtttggcg 2580
73 gtttttggtg atattggtaa tgcattatgat aaaggcttta ctaatgatac caaaattggt 2640
74 gcagggtgctg gtgttcgctg ggcacacact gtcgggtcaag ttcgtgttga tgtggcaact 2700
75 ggtgtcaaaag aagagggcaa tcccattaag ctgcattttt ttattggcac accattttaa 2760

```

77 <210> SEQ ID NO: 2

78 <211> LENGTH: 919

79 <212> TYPE: PRT

80 <213> ORGANISM: Moraxella catarrhalis

82 <400> SEQUENCE: 2

```

83 Met Ser Lys Pro Val Leu Phe Ala Asn Arg Ser Phe Met Pro Val Ala
84 1 5 10 15
85 Leu Ala Ala Tyr Leu Pro Leu Met Thr Ser Gln Ala Leu Ala Gln Gln
86 20 25 30
87 Asn Asn Pro Ala Asn Ile Ile Asn His Val Pro Ala His Asp Thr Ala
88 35 40 45
89 Ile Asn Gln Ala Lys Ala Gly Asn Pro Pro Val Leu Leu Thr Pro Glu
90 50 55 60
91 Gln Ile Gln Ala Arg Leu Asn Ala Ala Gly Leu Asn Ala Lys Pro Gln
92 65 70 75 80
93 Ser Gln Ala Leu Asp Val Val Asn Phe Asp Asp Gln Ser Pro Ile Ser
94 85 90 95
95 Arg Ile Gly Glu Gln Ser Pro Pro Leu Gly Leu Asp Met Ser Val Ile
96 100 105 110
97 Glu Glu Thr Thr Pro Leu Ser Leu Glu Glu Leu Phe Ala Gln Glu Ser
98 115 120 125
99 Thr Glu Met Gly Ile Asn Pro Asn Asp Tyr Ile Pro Glu Tyr Gln Gly
100 130 135 140
101 Glu Gln Pro Asn Ser Glu Val Val Val Pro Pro Thr Leu Glu Pro Glu
102 145 150 155 160
103 Lys Pro Gly Leu Ile Lys Arg Leu Tyr Ala Arg Leu Phe Asn Asp Gly
104 165 170 175
105 Val Asn Lys Val Pro Arg Leu Lys Ala Lys Phe Tyr Gln Ser Ser Gln
106 180 185 190
107 Ser Gly Glu Thr Ser Ala Ile Gly Ser Ser His Gln Lys Thr Glu Pro

```

RAW SEQUENCE LISTING

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

```

108          195          200          205
109 Tyr Ala Asn Ile Lys Ala Ala Leu Glu Asp Ile Thr Gln Glu Ser Ala
110          210          215          220
111 Met Asp Leu Asn Gly Ser Ile Pro Arg Leu Arg Gln Thr Ala Leu Val
112 225          230          235          240
113 Ala Ala Arg Ala Val Gly Tyr Tyr Asp Ile Asp Leu Ser Ile Ile Arg
114          245          250          255
115 Asn Ser Ile Gly Glu Val Asp Val Ile Ile His Asp Leu Gly Glu Pro
116          260          265          270
117 Val Tyr Ile Asp Tyr Arg Ala Val Glu Val Arg Gly Glu Gly Ala Asp
118          275          280          285
119 Asp Lys Ala Phe Thr Thr Val Ala Asp Glu Val Pro Leu Leu Ile Gly
120          290          295          300
121 Asp Val Phe His His Gly Lys Tyr Glu Thr Lys Lys Asn Leu Ile Glu
122 305          310          315          320
123 Asn Ala Ser Ala Glu His Gly Tyr Phe Asp Gly Arg Trp Leu Asp Arg
124          325          330          335
125 Ser Val Asp Val Ile Leu Pro Asp Asn Thr Ala Asp Val Ser Leu Ile
126          340          345          350
127 Tyr Asp Thr Gly Thr Gln Tyr Arg Phe Asp Glu Val Val Phe Phe Thr
128          355          360          365
129 Ile Asp Pro Lys Thr Asn Gln Leu Thr Thr Asp Pro Asp Lys Leu Pro
130          370          375          380
131 Val Lys Arg Glu Leu Leu Glu Gln Leu Leu Thr Val Asn Met Gly Glu
132 385          390          395          400
133 Ala Tyr Asn Leu Gln Ala Val Arg Ala Leu Ser Asn Asp Leu Ile Ala
134          405          410          415
135 Thr Arg Tyr Phe Asn Met Val Asn Thr Glu Ile Val Phe Pro Glu Arg
136          420          425          430
137 Glu Gln Ile Gln Asn Asp Gln Val Ser Phe Glu Gln Ser Ser Ser Ser
138          435          440          445
139 Arg Thr Glu Pro Ala Gln Val Asp Glu Ser Thr Leu Glu Pro Val Ile
140          450          455          460
141 Glu Thr Val Glu Leu Thr Asp Gly Ile Leu Met Asp Ile Ser Pro Ile
142 465          470          475          480
143 Glu Phe Ser Ala Ser Asn Leu Ile Gln Asp Lys Leu Asn Leu Val Ala
144          485          490          495
145 Ala Lys Ala Arg His Leu Tyr Asp Met Pro Asp Asp Arg Val Leu Ala
146          500          505          510
147 Ile Asn His Asp Asp Gly Val Asn Arg Ser Ile Leu Gly Arg Ile Ser
148          515          520          525
149 Asp Ala Val Ser Ala Val Ala Arg Ala Ile Leu Pro Asp Glu Ser Glu
150          530          535          540
151 Asn Glu Val Ile Asp Leu Pro Glu Arg Thr Ala Leu Ala Asn Arg Lys
152 545          550          555          560
153 Thr Pro Ala Asp Val Tyr Gln Ser Lys Lys Val Pro Leu Tyr Val Phe
154          565          570          575
155 Val Ala Ser Asp Lys Pro Arg Asp Gly Gln Ile Gly Leu Gly Trp Gly
156          580          585          590

```

RAW SEQUENCE LISTING

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

```

157 Ser Asp Thr Gly Thr Arg Leu Val Thr Lys Phe Glu His Asn Leu Ile
158          595                      600                      605
159 Asn Arg Asp Gly Tyr Gln Ala Gly Ala Glu Leu Arg Leu Ser Glu Asp
160          610                      615                      620
161 Lys Lys Gly Val Lys Leu Tyr Ala Thr Lys Pro Leu Ser His Pro Leu
162 625          630                      635                      640
163 Asn Asp Gln Leu Arg Ala Thr Leu Gly Tyr Gln Gln Glu Val Phe Gly
164          645                      650                      655
165 His Ser Thr Asn Gly Phe Asp Leu Ser Thr Arg Thr Leu Glu His Glu
166          660                      665                      670
167 Ile Ser Arg Ser Ile Ile Gln Asn Gly Gly Trp Asn Arg Thr Tyr Ser
168          675                      680                      685
169 Leu Arg Tyr Arg Leu Asp Lys Leu Lys Thr Gln Ala Pro Pro Glu Thr
170          690                      695                      700
171 Trp Gln Asp Leu Pro Val Asp Phe Val Asn Gly Lys Pro Ser Gln Glu
172 705          710                      715                      720
173 Ala Leu Leu Ala Gly Val Ala Val His Lys Thr Val Ala Asp Asn Leu
174          725                      730                      735
175 Val Asn Pro Met Arg Gly Tyr Arg Gln Arg Tyr Ser Leu Glu Val Gly
176          740                      745                      750
177 Ser Ser Gly Leu Val Ser Asp Ala Asn Met Ala Ile Ala Arg Ala Gly
178          755                      760                      765
179 Ile Ser Gly Val Tyr Ser Phe Gly Asp Asn Ala Tyr Gly Ser Asn Arg
180          770                      775                      780
181 Ala His Gln Met Thr Gly Gly Ile Gln Ala Gly Tyr Ile Trp Ser Asp
182 785          790                      795                      800
183 Asn Phe Asn His Val Pro Tyr Arg Leu Arg Phe Phe Ala Gly Gly Asp
184          805                      810                      815
185 Gln Ser Ile Arg Gly Tyr Ala His Asp Ser Leu Ser Pro Ile Ser Asp
186          820                      825                      830
187 Lys Gly Tyr Leu Thr Gly Gly Gln Val Leu Ala Val Gly Thr Ala Glu
188          835                      840                      845
189 Tyr Asn Tyr Glu Phe Met Lys Asp Leu Arg Leu Ala Val Phe Gly Asp
190          850                      855                      860
191 Ile Gly Asn Ala Tyr Asp Lys Gly Phe Thr Asn Asp Thr Lys Ile Gly
192 865          870                      875                      880
193 Ala Gly Val Gly Val Arg Trp Ala Ser Pro Val Gly Gln Val Arg Val
194          885                      890                      895
195 Asp Val Ala Thr Gly Val Lys Glu Glu Gly Asn Pro Ile Lys Leu His
196          900                      905                      910
197 Phe Phe Ile Gly Thr Pro Phe
198          915
200 <210> SEQ ID NO: 3
201 <211> LENGTH: 2670
202 <212> TYPE: DNA
203 <213> ORGANISM: Moraxella catarrhalis
205 <400> SEQUENCE: 3
206 caacaaaata accctgcaaa catcatcaat catgtaccgc ctcacgacac cgccatcaat      60
207 caagcaaagg caggcaatcc gcctgttttg ctaacacctg agcagataca agcacgcctt      120

```

RAW SEQUENCE LISTING

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

```

208 aatgctgctg gactgaatgc taagcccca tcacaagctt tggatgttgt caattttgat 180
209 gatcaatcgc cgatatctcg tatcgggtgag caatcacccc ctttgggttt ggatatgtcg 240
210 gtcatcgaag aaaccacacc gctaagcttg gaggaattat ttgctcaaga atctactgag 300
211 atgggaatca atccaaatga ttatatcca gaatatcaag gcgagcaacc taatagttag 360
212 gtggttgtag caccgacatt agaacctgaa aaaccagggt tgatcaagcg tctttatgca 420
213 cgctatttta atgatggtgt caataagggt cctaggctta aggcaaaatt ttatcaatca 480
214 tcgcaatcag gcgaaaccag tgcgattggg tcatcgcac aaaaaacaga gccttatgca 540
215 aatatcaaag cagcacttga agacatcacc caagagtcag cgatggattt gaatggctct 600
216 atcccacgcc taaggcaaac tgctttggtg gcagcgcgtg ctgtcggtta ttatgatatt 660
217 gatttatcaa tcataagaaa tagcatcggg gaggtggatg tcatcatcca tgatttaggt 720
218 gaacctgttt atattgatta tcgagcgggt gaggtacgag gtgaagggtg tgatgataaa 780
219 gcatttacta ccgtggcgga tgaggtgcca ttgctgatcg gcgatgtctt tcatcacggc 840
220 aagtacgaaa ccaaaaaaaaa tctcatcgaa aatgccagtg ctgaacatgg atattttgat 900
221 gggcggttggc tggatcgttc agttgatgta attttgccag ataataccgc tgatgtcagc 960
222 ttaatttatg atacaggtag gcagtatcgc tttgatgagg tgataatttt taccattgat 1020
223 cctaaaacca atcaattgac aaccgatcca gataagctgc cagttaaacg agaattactt 1080
224 gagcagttac tcaccgttaa catgggagag gcttacaatt tacaggcggt gcgtgcactt 1140
225 tcaaatgatt tgattgccac acggtatttt aatatggtga ataccgagat tgtctttcca 1200
226 gagcgtgaac agatccaaaa cgaccaagtg agctttgagc agtcttcaag tagccgtact 1260
227 gaaccagcac aagttgatga aagcacactt gaacctgtca ttgaaaccgt tgagctaacg 1320
228 gatgggatat taatggatat ttcgcccatc gaatttagtg catctaactt gattcaagac 1380
229 aagctaaatt tgggtggctg caaggctcgc catttatatg acatgcctga tgataggggtg 1440
230 cttgccatca atcatgatga tggcgtaaat cgctctattt tgggcagaat cagcgatgcc 1500
231 gtatctgccg ttgcacgtgc tattttacct gatgaatctg aaaatgaggt aatagatttg 1560
232 cccgagcgta ccgcattggc taatcgcaag acccctgctg atgtctatca aagtaaaaaa 1620
233 gtgccgctat atgtctttgt ggcgagtgat aaaccccgag atggtcaa at tggtttgggc 1680
234 tggggatcgg acacaggtac ccgcctagtc acaaaatttg agcataattt gattaatcgt 1740
235 gatggctatc aagcagcgcg tgagctaaga ctgtctgagg ataaaaaagg ggtcaagtta 1800
236 tatgccacca aaccgcttag ccacctcta aatgatcagc taagagcaac tttgggttat 1860
237 caacaagaag tttttggtca ctctaccaat ggttttgatt tatccacacg caccctagag 1920
238 catgagatta gccgcagtat tatccaaaat ggtggctgga atcgtaacta ttcattgcgt 1980
239 tatcgtcttg ataagcttaa aacccaagca cccctgaaa catggcagga tttaccagt 2040
240 gattttgtca atggtaaagg aagccaagag gcgttatttg cagggtgtgc tgtgcataaa 2100
241 acggttgtag ataatttggg taatccgatg cgtggctatc gtcagcgata ttccttagag 2160
242 gttggctcaa gcggtttggg atcggtatgt aatatggcta ttgctcgagc tggatttagt 2220
243 ggcgtgtata gttttgggga taatgcttat ggcagcaatc gtgcccata gatgactggt 2280
244 ggcatacaag caggatacat ttggtcggat aattttaatc atgtgccata tcgtttgcgt 2340
245 ttttttgctg gtggcgacca aagtattcgt ggatatgcac atgacagttt atcacctata 2400
246 tcagataagg gttatctgac aggcgggtcaa gtattggcgg ttggtacagc tgaatataat 2460
247 tatgaattta tgaaagattt gcggtttggc gtttttggtg atattggtaa tgcttatgat 2520
248 aaaggcttta ctaatgatac caaaattggg gcagggtgct gtgttcgctg ggcacacct 2580
249 gtcggtcaag ttcgtgttga tgtggcaact ggtgtcaaa aagagggcaa tcccatataa 2640
250 ctgcattttt ttattggcac accattttta 2670
252 <210> SEQ ID NO: 4
253 <211> LENGTH: 889
254 <212> TYPE: PRT
255 <213> ORGANISM: Moraxella catarrhalis
257 <400> SEQUENCE: 4
258 Gln Gln Asn Asn Pro Ala Asn Ile Ile Asn His Val Pro Ala His Asp

```

VERIFICATION SUMMARY

DATE: 03/13/2002

PATENT APPLICATION: US/09/914,168

TIME: 13:41:31

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\03132002\I914168.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date